

Sanov and Central Limit Theorems for output statistics of quantum Markov chains

Merlijn van Horssen^{1,*} and Mădălin Guță^{2,†}

¹*School of Physics and Astronomy,
University of Nottingham, Nottingham, NG7 2RD, UK*

²*School of Mathematical Sciences,
University of Nottingham, Nottingham, NG7 2RD, UK*

(Dated: 22nd September 2014)

Abstract

In this paper we consider the statistics of repeated measurements on the output of a quantum Markov chain. We establish a large deviations result analogous to Sanov's theorem for the empirical measure associated to finite sequences of consecutive outcomes of a classical stochastic process. Our result relies on the construction of an extended quantum transition operator (which keeps track of previous outcomes) in terms of which we compute moment generating functions, and whose spectral radius is related to the large deviations rate function. As a corollary to this we obtain a central limit theorem for the empirical measure. Such higher level statistics may be used to uncover critical behaviour such as dynamical phase transitions, which are not captured by lower level statistics such as the sample mean. As a step in this direction we give an example of a finite system whose level-one rate function is independent of a model parameter while the level-two rate is not.

* E-mail: merlijn.vanhorssen@nottingham.ac.uk

† E-mail: madalin.guta@nottingham.ac.uk

I. INTRODUCTION

Quantum Markov processes are effective mathematical models for describing a wide class of open quantum dynamics where the environment interacts weakly with the system [1–3]. In the discrete time version, a quantum Markov chain consists of a system which interacts successively with identically prepared ancillas (input), via a fixed unitary transformation, cf. Figure 1. After the interaction, the ancillas (output) are in a finitely correlated state [4] which carries information about the dynamics. Such information can be extracted by performing successive measurements on the outgoing ancillas. Similar models are used in continuous time in the input-output formalism of quantum optics [1], where output measurements (e.g. photon counting or homodyne) are used to monitor and control the system [5]. Interest in such systems has grown with recent experimental progress in quantum optics and open quantum many-body systems [6–9], in particular from the point of view of dynamical phase transitions [10–19] and system identification [20, 21].

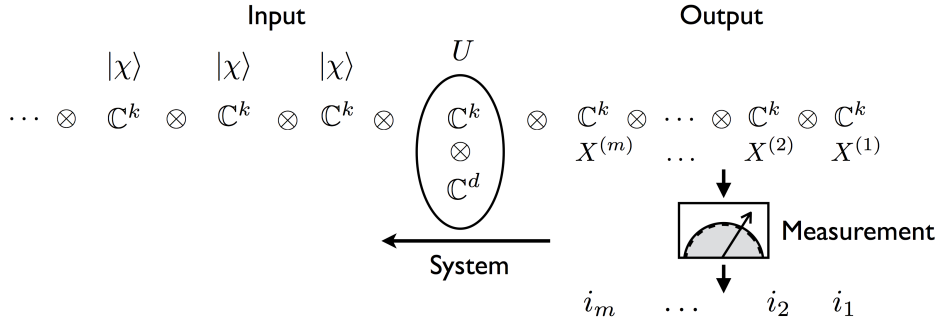


Figure 1: A quantum Markov chain consists of a system \mathbb{C}^d interacting successively with a sequence of identically prepared input ancillas with space \mathbb{C}^k via the unitary operator U . Performing a measurement of a single-site operator X on each of the output sites results in a sequence of measurement outcomes i_1, i_2, \dots, i_m .

Dynamical phase transitions in open quantum systems are visible through spectral properties of the generator of the dynamics [12], and through the thermodynamics of jump trajectories [22–25]; the large deviations approach to quantum phase transitions exploits both these features to uncover dynamical phases [17, 26]. In this article we establish a large deviation principle (LDP) for *counts of sequences of successive outcomes* in the output of a quantum Markov chain, which we refer to as the Sanov theorem for such systems. Sanov theorems have been considered in quantum systems before, both generally [27] and in the

context of quantum hypothesis testing [28–31]; our result extends previous work on LDPs for counting statistics, by considering counts of sequences of outcomes on an arbitrary number of subsequent sites.

The Sanov theorem for classical Markov chains [32, 33] establishes an LDP for the empirical measure (the proportion of states visited by the chain) and the pair empirical measure (the frequency with which jumps between pairs of states occur) (see Fig. 2c); a result used for example in [34] to characterise phase transitions in the Curie-Weiss-Potts model. In non-equilibrium statistical mechanics, a dynamical phase transition may occur when the Gärtner-Ellis theorem for some statistic fails to hold [34, 35], corresponding to a non-analyticity of the associated LD rate function.

To establish the LDP for the higher-level statistics in the Sanov theorem, we express the associated sequence of moment generating functions in terms of an extended transition operator, constructed from the original quantum Markov chain transition operator (analogous to the method in [27]). The LDP is then obtained via the Gärtner-Ellis theorem; the corresponding LD rate function is obtained in terms of the spectral radius of a perturbation of this new transition operator, where primitivity of the original transition operator ensures that this rate function is analytic.

In general, the existence of an LDP for a stochastic process does not imply that the process also satisfies a central limit theorem. However, in certain cases [36] the proof of the LDP may be extended to produce a central limit theorem; as a corollary to our main result, we thus state a central limit theorem for the empirical measures.

The paper is organised as follows: in Sec. II we introduce the background to our result. We briefly review the theory of large deviations in Sec. II A, we introduce quantum Markov chains in Sec. II B and consider measurements on the output of a quantum Markov chain in Sec. II C. In Sec. III we state our main result, Thm. 3, and Corollary 1 (preceded by definitions and results directly related to our main result). To illustrate our result we discuss two examples in Sec. IV.

II. BACKGROUND

In this section we present a brief review of the basic concepts of large deviations needed in this paper, and introduce the set-up of quantum Markov chains. For good introductions

to the theory and applications of large deviations we refer to the monographs [32–34, 37].

A. Large deviations

Let Y_1, Y_2, \dots be a sequence of independent and identically distributed (i.i.d.) \mathbb{R}^d -valued random variables defined on a probability space $(\Omega, \Sigma, \mathbb{P})$. The law of large numbers (LLN) states that if $\mathbb{E}(Y_i) = y$ exists, then the average

$$X_n := \frac{1}{n} \sum_{i=1}^n Y_i$$

converges to y almost surely as $n \rightarrow \infty$. The Central Limit Theorem (CLT) characterises the speed of convergence, and shows that the fluctuations around the mean decrease as $n^{-1/2}$, and are asymptotically normally distributed:

$$\sqrt{n}(X_n - y) \xrightarrow{\mathcal{L}} N(0, v).$$

Here \mathcal{L} denotes the convergence in law (distribution) and $N(0, v)$ is the normal distribution with mean zero and variance $v := \mathbb{E}(Y_i^2) - \mathbb{E}(Y_i)^2$. As the distribution of X_n concentrates around y , one would like to know how the probability of staying away from y decreases with n . Such large deviations have exponentially small probabilities i.e.

$$\mathbb{P}(X_n \geq y + a) \sim \exp(-nI(a)) \tag{1}$$

in a sense which will be made precise below.

In another example, suppose Y_1, Y_2, \dots are i.i.d. random variables with values in $\{1, \dots, d\}$ and common probability distribution $p_i = \mathbb{P}(Y = i)$. The *empirical measure* $\hat{\mathbb{P}}_n$ is defined as the random probability distribution given by the frequencies with which the different values in $\{1, \dots, d\}$ occur (where the indicator function is defined by $1_i(Y) = 1$ if $Y = i$, and 0 otherwise)

$$\hat{\mathbb{P}}_n(i) = \frac{1}{n} \sum_{j=1}^n 1_i(Y_j). \tag{2}$$

Again, by the LLN we have $\hat{\mathbb{P}}_n(i) \rightarrow p_i$, and the fluctuations around the mean can be described by the CLT. The large deviations are characterised by the following result known as the Sanov theorem. If R is a measurable subset of the simplex of probability distributions, and $p \notin R$, then

$$\mathbb{P}(\hat{\mathbb{P}}_n \in R) \sim \exp\left(-n \inf_{q \in R} I(q|p)\right)$$

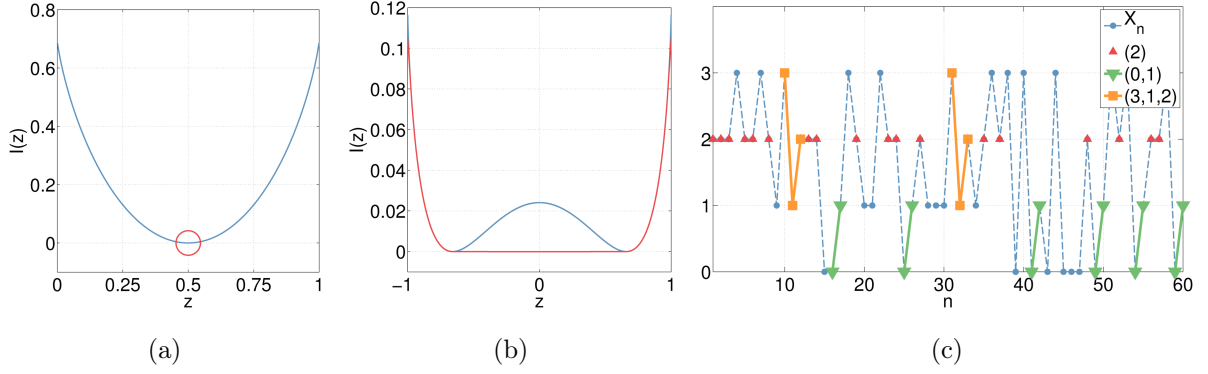


Figure 2: Typical large deviations rate functions. **(a)**: rate function $I(z)$ on the interval $[0, 1]$ associated to the sample mean of a fair coin toss, with minimum of I indicated at $z = 1/2$; **(b)**: example of a non-convex rate function (blue) and its convex envelope (red) obtained from the Gärtner-Ellis theorem. **(c)**: Sample trajectory X_n with events associated to different levels, e.g. visits to state 2, jumps from 0 to 1, etc., which are used to compute the empirical measures.

where the rate function $I(q|p)$ is the relative entropy

$$I(q|p) = \sum_{i=1}^d p_i \log(p_i/q_i).$$

We summarise (2) by saying that "a large deviation event will happen in the least unlikely of all the unlikely ways" [32].

We can now describe a more general set-up of large deviations theory [32–34, 37], and formulate one of the key mathematical tools used later in our paper. A sequence $\{\mu_n : n \in \mathbb{N}\}$ probability distributions on \mathbb{R}^d is said to satisfy a *large deviation principle* (LDP) if there exists a lower semicontinuous function $I : \mathbb{R}^d \rightarrow [0, \infty]$, called a *rate function*, such that for all measurable subsets $B \subset \mathbb{R}^d$

$$\begin{aligned} - \inf_{x \in B^0} I(x) &\leq \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(B) \\ &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mu_n(B) \leq - \inf_{x \in \bar{B}} I(x). \end{aligned}$$

Here B^0 and \bar{B} denote the interior and closure of B , respectively; if these sets coincide, the LDP may be expressed in the intuitive form used above $\mu_n(B) \sim \exp(-n \inf_{x \in B} I(x))$.

In our results we will employ a theorem due to Gärtner [38] and Ellis [39], which provides a necessary condition for obtaining an LDP; we state it here with stronger assumptions,

equivalent to the result in [38].

Theorem 1. Let $\{\Gamma_n : n \in \mathbb{N}\}$ be the moment generating function associated to μ_n

$$\Gamma_n(t) = \int_{\mathbb{R}^d} e^{n\langle t, x \rangle} d\mu_n(x), \quad t \in \mathbb{R}^d,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product. Suppose the limit

$$F(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \Gamma_n(t) \quad (3)$$

is finite for all $t \in \mathbb{R}^d$ and $F : \mathbb{R}^d \rightarrow \mathbb{R}$ is a differentiable function. Then $\{\mu_n : n \in \mathbb{N}\}$ satisfies an LDP with rate function I given by the Legendre-Fenchel transform of F ,

$$I(x) = \sup_{t \in \mathbb{R}^d} \{\langle t, x \rangle - F(t)\}. \quad (4)$$

The Gartner-Ellis theorem implies that the rate function $I(\cdot)$ in (1) is given by the Legendre transform of $\Gamma(t) = \mathbb{E}(e^{tY})$, and can also be employed for proving the Sanov theorem, where μ_n is the distribution of the empirical measure $\hat{\mathbb{P}}_n$ seen as a vector in \mathbb{R}^d .

The Sanov Theorem can be extended [40–42] to empirical measures associated to an *irreducible*¹ Markov chain $\{X_n : n \in \mathbb{N}\}$ over a discrete state space $\{1, \dots, d\}$ with transition matrix Π . For instance, the empirical measure $\hat{\mathbb{P}}^{(1)}(i) := \sum_{j=1}^n 1_i(X_j)$ keeps track of the empirical frequencies associated to each state which by ergodicity converge the stationary distribution of the chain. This empirical measure satisfies an LDP on \mathbb{R}^d , where the rate function is given by the Legendre transform of the function $\lambda \mapsto \log r(\Pi_\lambda)$. Here $r(\cdot)$ denotes the *spectral radius*; the matrix Π_λ is a certain analytic perturbation of Π . Similarly, the pair-empirical measure

$$\hat{\mathbb{P}}^{(2)}(k, l) := \sum_{j=1}^{n-1} 1_{k,l}(X_j, X_{j+1}),$$

which encodes additional information about how the chain jumps from one state to another, also satisfies an LDP. These empirical measures of different orders are illustrated in Fig.2c on a sample trajectory of a Markov chain with four states. We note that this approach used to establish LDPs for the empirical measures associated to Markov chains bears similarities to the proof of the main result of this paper which deals with the output process of a quantum Markov chains.

¹ A Markov chain with transition matrix Π is called irreducible if for all $1 \leq i, j \leq d$ there is an integer n such that $\Pi^n(i, j) > 0$

B. Quantum Markov chains

A quantum Markov chain [3] consists of a system, or ‘memory’ with Hilbert space \mathbb{C}^d which interacts successively (moving from right to left) with a chain of identically prepared ancillas, or ‘noise units’ \mathbb{C}^k , via a unitary $U : \mathbb{C}^d \otimes \mathbb{C}^k \rightarrow \mathbb{C}^d \otimes \mathbb{C}^k$, cf. Figure 1. Physically, this may be seen a discrete-time model for the evolution of an open quantum system coupled to the environment in the Markov approximation, as shown in [43, 44].

We assume that the initial state of the noise units (or input) is a fixed vector $|\chi\rangle \in \mathbb{C}^d$, such that the output state is determined by the isometry

$$\begin{aligned} V : \mathbb{C}^d &\rightarrow \mathbb{C}^d \otimes \mathbb{C}^k \\ V : |\psi\rangle &\mapsto U(|\psi\rangle \otimes |\chi\rangle) = \sum_{i=1}^k V_i |\psi\rangle \otimes |i\rangle, \end{aligned} \quad (5)$$

where $\{|1\rangle, \dots, |k\rangle\}$ are the vectors of an orthonormal basis in \mathbb{C}^k . The operators $V_i = \langle i|U|\chi\rangle$ will be called the *Kraus operators* associated to the quantum Markov chain, and satisfy the normalisation condition $\sum_{i=1}^k V_i^* V_i = \mathbf{1}$. If the system starts in the state $|\psi\rangle$, we can apply (5) successively to express the *joint* state of n output units and the system as a matrix product state [45, 46]

$$|\psi^{(n)}\rangle = \sum_{i_1, \dots, i_n=1}^k V_{i_n} \cdots V_{i_1} |\psi\rangle |i_n, \dots, i_1\rangle \quad (6)$$

reflecting the inherent Markovian character of the dynamics. By tracing over the noise units we find that the reduced system dynamics is given by the semigroup

$$T_*^n : M_d \rightarrow M_d, \quad n \in \mathbb{N},$$

where

$$T_* : \rho \mapsto \sum_{i=1}^k V_i \rho V_i^* = \text{Tr}_{\mathbb{C}^k} (U \rho \otimes |\chi\rangle\langle\chi| U^*)$$

is a trace preserving completely positive map describing the system’s transition operator. For later purposes we note that the corresponding map in the Heisenberg representation is given by

$$\begin{aligned} T : M_d &\rightarrow M_d \\ T : A &\mapsto \sum_{i=1}^k V_i^* A V_i. \end{aligned}$$

Now, suppose that after the interaction we perform a projective measurement on each of the output noise units, with respect to the basis $\{|1\rangle, \dots, |k\rangle\}$. If $X^{(i)}$ denotes the outcome of the measurement of the i th unit, then the joint probability distribution of the measurement process is

$$p(i_1, \dots, i_n) = \mathbb{P}(X^{(1)} = i_1, \dots, X^{(n)} = i_n) = \|V_{i_n} \cdots V_{i_1} |\psi\rangle\|^2.$$

This process is not necessarily stationary but becomes so in the large n limit if T satisfies a certain ergodic property discussed in Sec. III below.

C. Empirical measures associated to the measurement process

Our main goal is to establish a large deviation principle for the empirical measure associated to chains of subsequent outcomes occurring in the measurement trajectory. For each $m \in \mathbb{N}$ we define the empirical measure $\hat{\mathbb{P}}_n^{(m)}$ over $\{1, \dots, k\}^m$ by

$$\hat{\mathbb{P}}_n^{(m)}(i_1, \dots, i_m) := \frac{1}{n - m + 1} \sum_{j=1}^{n-m+1} 1_{(i_1, \dots, i_m)}(X^{(j)}, \dots, X^{(j+m-1)}). \quad (7)$$

In the case $m = 1$ this keeps track of the frequencies of the different outcomes in a measurement trajectory, and the associated LDP has been established in [27]. A similar result holds in continuous time, for total counts statistics when a counting measurement is performed on the output, and this plays a key role in the theory of *dynamical phase transitions* [12, 24]. For $m = 2$ the empirical measure captures the statistics of pairs of subsequent outcomes. As we will show in Sec. IV, the higher level ($m > 1$) LD rates capture more information about the measurement process than the total counts, and therefore could be the basis of a more in-depth understanding of the theory of dynamical phase transitions.

The main tool in establishing the LDP will be the Gärtner-Ellis theorem 1. For this we need to compute the moment generating function

$$\Gamma_n^{(m)}(t) := \mathbb{E} \left(\exp(n \langle \hat{\mathbb{P}}_n^{(m)}, t \rangle) \right) \quad (8)$$

where $t \in \mathbb{R}^{k^m}$ and

$$\langle \hat{\mathbb{P}}_n^{(m)}, t \rangle = \sum_{i_1, \dots, i_m} t_{i_1, \dots, i_m} \hat{\mathbb{P}}_n^{(m)}(i_1, \dots, i_m).$$

We will show that the MGF can be expressed in terms of a certain extended transition operator which acts on the system but also takes into account the results of $m - 1$ measurement

outcomes. For this we consider the space

$$\mathcal{D}_d^{m,k} := M_d \otimes (\mathbb{C}^k)^{\otimes(m-1)},$$

seen as a block diagonal algebra with $k(m-1)$ blocks isomorphic to M_d . We will represent an element $Y \in \mathcal{D}_d^{m,k}$ by the diagonal elements $[Y]_{i_1, \dots, i_{m-1}} \in M_d$, where $i_1, \dots, i_{m-1} \in \{1, \dots, k\}$.

Lemma 1. Let $T_{t,m} : \mathcal{D}_d^{m,k} \rightarrow \mathcal{D}_d^{m,k}$ be the completely positive map defined as

$$[T_{t,m}(Y)]_{i_1, \dots, i_{m-1}} = \sum_{i_m=1}^k e^{t_{i_1, \dots, i_m}} V_{i_1}^* [Y]_{i_2, \dots, i_m} V_{i_1} \quad (9)$$

Then the moment generating function (8) can be expressed as

$$\Gamma_n^{(m)}(t) = \langle \psi | \sum_{i_1, \dots, i_{m-1}=1}^k [T_{t,m}^{n-m+1}(M^{(m)})]_{i_1, \dots, i_{m-1}} | \psi \rangle$$

where

$$[M^{(m)}]_{i_1, \dots, i_{m-1}} = V_{i_1}^* \cdots V_{i_{m-1}}^* V_{i_{m-1}} \cdots V_{i_1}.$$

In particular, $M^{(m)}$ is an eigenvector of $T_{0,m}$ with eigenvalue 1.

Proof. We denote by $X^{(l,m)}$ the \mathbb{R}^{k^m} -valued random variable which represents the outcomes on m subsequent sites: for a sequence of outcomes $\mathbf{i} = (i_1, \dots, i_m)$ we have

$$[X^{(l,m)} = \mathbf{i}] \Leftrightarrow [X^{(l)} = i_1, \dots, X^{(l+m-1)} = i_m]$$

where $X^{(j)}$ is the random variable associated to measurement outcomes of the observable X on site j . The moment generating function is then

$$\Gamma_n^{(m)}(t) = \mathbb{E} \left[\exp \left\langle t, \sum_{j=1}^{n-m+1} X^{(j,m)} \right\rangle \right].$$

Writing $P_{\mathbf{i}}^{(l,m)} := P_{i_1}^{(l)} \otimes \cdots \otimes P_{i_m}^{(l+m-1)}$ for the one dimensional projection on m subsequent sites associated to a sequence of outcomes $X^{(l,m)} = \mathbf{i}$, we have

$$\Gamma_n^{(m)}(t) = \left\langle \exp \left(\sum_{l=1}^{n-m+1} \sum_{\mathbf{i}} t_{\mathbf{i}} P_{\mathbf{i}}^{(l,m)} \right) \right\rangle,$$

where the expectation is taken with respect to the system-output state $|\psi^{(n)}\rangle$ defined in (6).

This gives

$$\Gamma_n^{(m)}(t) = \sum_{i_1, \dots, i_n=1}^k \langle \psi | V_{i_1}^* \cdots V_{i_n}^* V_{i_n} \cdots V_{i_1} | \psi \rangle \cdot \exp \left(\sum_{l=1}^{n-m+1} t_{i_l, \dots, i_{l+m-1}} \right). \quad (10)$$

Finally, a short computation shows that $\Gamma_n^{(m)}$ can be expressed as

$$\Gamma_n^{(m)}(t) = \langle \psi | \sum_{i_1, \dots, i_{m-1}=1}^k [T_{t,m}^{n-m+1} (M^{(m)})]_{i_1, \dots, i_{m-1}} | \psi \rangle$$

where

$$[M^{(m)}]_{i_1, \dots, i_{m-1}} = V_{i_1}^* \cdots V_{i_{m-1}}^* V_{i_{m-1}} \cdots V_{i_1}.$$

The eigevale property

$$T_{0,m}(M^{(m)}) = M^{(m)}$$

follows directly from the definition of $T_{t,m}$ and the normalisation $\sum_i V_i^* V_i = \mathbf{1}$.

□

III. MAIN RESULT

In this section we recall notions of irreducibility and primitivity (Def. 1), and state existing results (Thm. 2 and Lem. 2) concerning irreducible maps. We then state and prove our main result, the Sanov theorem for the empirical measure (7) in Thm. 3. We note that the following results are all in the context of positive maps on finite-dimensional C^* -algebras. In our theorem, we are considering positive linear maps on algebras with a block form $\bigoplus_j M_{d_j}$, where each M_{d_j} is an algebra of $d_j \times d_j$ matrices with complex entries; algebras of this form are a particular class of C^* -algebras [47].

Definition 1 ([48]). Let \mathcal{A} be a finite-dimensional C^* -algebra and let R be a positive linear map on \mathcal{A} . Then R is called

- (i) *irreducible* if there exists $n \in \mathbb{N}$ such that $(\text{Id} + R)^n$ is strictly positive, i.e. $(\text{Id} + R)^n(A) > 0$ for all positive operators A , and
- (ii) *primitive* if there exists an $n \in \mathbb{N}$, such that R^n is strictly positive.

Primitivity is a stronger requirement than irreducibility. The following theorem (see [48–50]) collects the essential properties of irreducible (primitive) quantum transition operators needed in this paper.

Theorem 2 (quantum Perron-Frobenius [48–50]). Let \mathcal{A} be a finite-dimensional C^* -algebra and let R be a positive linear map on \mathcal{A} . Denote by $r(R) := \max_i |\lambda_i|$ the *spectral radius* of R , where $\{\lambda_1, \dots, \lambda_{d^2}\}$ are the (complex) eigenvalues of R arranged in decreasing order of magnitude. Then

- (i) $r(R)$ is an eigenvalue of R , and it has a positive eigenvector.
- (ii) If additionally, R is unit preserving then $r(R) = 1$ with eigenvector $\mathbf{1}$.
- (iii) If additionally, R is irreducible then $r(R)$ is a nondegenerate eigenvalue for R and R_* , and both corresponding eigenvectors are strictly positive.
- (iv) If additionally, R is primitive then $|\lambda_i| < r(R)$ for all eigenvalues other than $r(R)$.

As a corollary, if the Markov transition operator T is irreducible then it has a unique full rank stationary state (i.e. $T_*(\rho_{\text{ss}}) = \rho_{\text{ss}}$), and if T is also primitive then any state converges to the stationary state in the long run (mixing or ergodicity property);

$$\lim_{n \rightarrow \infty} (T_*)^n(\rho) = \rho_{\text{ss}},$$

or in the Heisenberg picture

$$\lim_{n \rightarrow \infty} T^n(A) = \text{tr}[A\rho_{\text{ss}}]\mathbf{1}.$$

For reader's convenience we state here a lemma [27] which will be used in the proof of the main theorem.

Lemma 2 ([27]). Let \mathcal{A} be a finite-dimensional C^* -algebra, let R be a positive linear map on \mathcal{A} and suppose that R has a strictly positive eigenvector. Then for any state $\varphi : \mathcal{A} \rightarrow \mathbb{C}$ and any positive $X \in \mathcal{A}$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \varphi(R^n(X)) = \log r(R).$$

We now state and prove our main result.

Theorem 3 (Sanov theorem for quantum Markov chains). Consider a quantum Markov chain on M_d whose transition operator T is primitive. Then the level m empirical measure (7) satisfies a large deviations principle on \mathbb{R}^{k^m} . The rate function is the Legendre transform of $\log r(\tilde{T}_{t,m})$, where $r(\tilde{T}_{t,m})$ is the spectral radius of a certain restriction $\tilde{T}_{t,m}$ of the extended transition operator defined in (9).

Proof. By the Gartner-Ellis Theorem 1 it suffices to show that the limit

$$F^{(m)}(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \Gamma_n^{(m)}(t) \quad (11)$$

exists for all t and $F^{(m)}$ is a continuous function. As it is defined, $T_{t,m}$ may not satisfy the conditions of Lem. 2, but we will show that (11) holds the same when $T_{t,m}$ is replaced by a certain restriction $\tilde{T}_{t,m}$ which does satisfy the conditions.

1. Invariance. Let \mathcal{B}_m be the (non-unital) subalgebra of $\mathcal{D}_d^{m,k}$ given by

$$\mathcal{B}_m = \bigoplus_{i_1, \dots, i_{m-1}}^k Q_{i_1, \dots, i_{m-1}} M_d Q_{i_1, \dots, i_{m-1}}$$

where Q is the projection onto the support of $M^{(m)}$. We will show that \mathcal{B}_m is invariant under $T_{t,m}$, i.e.

$$[T_{t,m}(Y)]_{i_1, \dots, i_{m-1}} \in Q_{i_1, \dots, i_{m-1}} M_d Q_{i_1, \dots, i_{m-1}}$$

for every $Y \in \mathcal{B}_m$. For this it suffices to show that

$$V_{i_1}^* [Y]_{i_2, \dots, i_m} V_{i_1}(u) = 0$$

for all $u \in \ker [M^{(m)}]_{i_1, \dots, i_{m-1}}$, or equivalently (since $\ker(A^*A) = \ker(A)$) all $u \in \ker(V_{i_{m-1}} \cdots V_{i_1})$.

Now

$$[Y]_{i_2, \dots, i_m}(v) = 0 \quad \text{for all } v \in \ker(V_{i_m} \cdots V_{i_2})$$

and we either have $u \in \ker(V_{i_1})$ or $V_{i_1}(u) \in \ker(V_{i_{m-1}} \cdots V_{i_2})$ we conclude that for every $u \in \ker [Y]_{i_1, \dots, i_{m-1}}$ we have $u \in \ker([T(Y)]_{i_1, \dots, i_{m-1}})$, proving that $T_{t,m}$ leaves \mathcal{B}_m invariant.

Let us denote by $\tilde{T}_{t,m}$ the restriction of $T_{t,m}$ to \mathcal{B}_m . Then, since $M^{(m)} \in \mathcal{B}_m$ the moment generating function can be expressed as

$$\Gamma_n^{(m)}(t) = \langle \psi | \sum_{i_1, \dots, i_{m-1}=1}^k \left[\tilde{T}_{t,m}^{n-m+1} (M^{(m)}) \right]_{i_1, \dots, i_{m-1}} | \psi \rangle$$

2. Primitivity. Since $\tilde{T}_{t,m} > c\tilde{T}_{0,m}$ for some positive constant c , it suffices to show that there exists $n \in \mathbb{N}$ such that, for any $X \in \mathcal{B}_m$,

$$\tilde{T}_{0,m}^n(X) \geq c' \mathbf{1}_{\mathcal{B}}$$

for some $c' > 0$. We can assume that $n > m$, in which case

$$\begin{aligned} & \left[\tilde{T}_{0,m}^n(X) \right]_{i_1, \dots, i_{m-1}} \\ &= \sum_{i_m, \dots, i_{m+n-1}=1}^k V_{i_1}^* \cdots V_{i_n}^* [X]_{i_{n+1}, \dots, i_{m+n-1}} V_{i_n} \cdots V_{i_1} \\ &= V_{i_1}^* \cdots V_{i_{m-1}}^* \left(\sum_{i_m, \dots, i_n=1}^k V_{i_m}^* \cdots V_{i_n}^* \hat{X} V_{i_n} \cdots V_{i_m} \right) V_{i_{m-1}} \cdots V_{i_1} \end{aligned}$$

where \hat{X} is the sum of the blocks of X given by

$$\hat{X} = \sum_{i_1, \dots, i_{m-1}=1}^k [X]_{i_1, \dots, i_{m-1}}.$$

The remaining sum can be written in terms of the original transition operator T associated to the quantum Markov chain; recall that for $Y \in M_d$

$$T(Y) = \sum_{i=1}^k V_i^* Y V_i$$

from which we obtain the expression

$$\left[\tilde{T}_{0,m}^n(X) \right]_{i_1, \dots, i_m} = V_{i_1}^* \cdots V_{i_{m-1}}^* T^{n-m+1}(\hat{X}) V_{i_{m-1}} \cdots V_{i_1}.$$

Since the original Markov chain is assumed to be primitive, there exists $r \in \mathbb{N}$ such that $T^r(\hat{X}) \geq c \mathbf{1}$ for some $c > 0$. Therefore, with $n \geq r + m - 1$,

$$\begin{aligned} \left[\tilde{T}_{0,m}^n(X) \right]_{i_1, \dots, i_{m-1}} &\geq c V_{i_1}^* \cdots V_{i_{m-1}}^* V_{i_{m-1}} \cdots V_{i_1} \\ &\geq c' Q_{i_1, \dots, i_{m-1}} \\ &= c' [\mathbf{1}_{\mathcal{B}}]_{i_1, \dots, i_{m-1}}. \end{aligned}$$

Using the invariance and primitivity property, we can apply Lemma 2, to find that the limiting moment generating function is given by

$$F^{(m)}(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \Gamma_n^{(m)}(t) = r(\tilde{T}_{t,m}).$$

Moreover, since $\tilde{T}_{t,m}$ is an analytic perturbation in t of $\tilde{T}_{0,m}$, the spectral radius $t \mapsto r(T_{t,m})$ is a smooth function [51], so the large deviation principle follows from the Gärtner-Ellis theorem.

□

As a corollary to our main result, we establish a central limit theorem for each of the empirical measures.

Corollary 1. Let $\{\hat{\mathbb{P}}_n^{(m)}\}$ be the sequence of distributions of the empirical measure of length m defined in Eq. (7). Then $\{\hat{\mathbb{P}}_n^{(m)}\}$ satisfies the Central Limit Theorem: that is (cf. Eq. (2)), as $n \rightarrow \infty$,

$$\sqrt{n} \left(\hat{\mathbb{P}}_n^{(m)} - p^{(m)} \right) \xrightarrow{\mathcal{D}} N(0, V^{(m)})$$

where \mathcal{D} denotes convergence in distribution. Here $p^{(m)}$ and $V^{(m)}$ are the mean and variance with respective components

$$p_i^{(m)} = \left. \frac{\partial \log r(\tilde{T}_{t,m})}{\partial t_i} \right|_{t=0}, \quad V_{i,j}^{(m)} = \left. \frac{\partial^2 \log r(\tilde{T}_{t,m})}{\partial t_i \partial t_j} \right|_{t=0}, \quad 1 \leq i, j \leq k^m.$$

Proof. Our main result relied on the convergence of the logarithmic moment generating functions $\frac{1}{n} \log \Gamma_n^{(m)}(t)$ in Eq. (11) to $F^{(m)}(t)$. For the purpose of establishing an LDP, it was sufficient to consider only real values for the parameter t . However, the same analytic perturbation arguments can be used to prove that locally in complex neighbourhood of $t = 0$, the Eq. (11) holds and the limiting function $F^{(m)}(z)$ is analytic in that region. The Central Limit Theorem for $\{\hat{\mathbb{P}}_n^{(m)}\}$ is then a consequence of the main result of [36].

□

We note that since the sequence $\{\hat{\mathbb{P}}_n^{(m)}\}$ is effectively a sequence of probability distributions on the probability simplex of dimension $k^m - 1$, the variance $V^{(m)}$ is degenerate in that at least one of its eigenvalues vanishes, corresponding to the degree of freedom in \mathbb{R}^{k^m} orthogonal to the probability simplex.

We end this section with a brief note on what happens when the transition operator is not irreducible. The Gärtner-Ellis theorem relies on differentiability of the logarithmic moment generating function $F(t)$ defined in Eq. 3 (an hypothesis which may be weakened to smoothness of $F(t)$ in a neighbourhood of the origin). We use irreducibility to ensure

that the spectral radius, and therefore the logarithmic moment generating function, is a differentiable function.

If $F(t)$ is not differentiable at $t = 0$, the first moments obtained as $\partial_t F(t)$ at $t = 0$ are not well defined, corresponding to a breaking down of a law of large numbers. On the level of the transition operator this corresponds to the dynamics breaking up into two disjoint parts, each with its own logarithmic moment generating function. In this case an LDP may still hold, but with a non-convex rate function. The rate function obtained from the Gärtner-Ellis theorem is a Legendre transform, and is the convex envelope of the actual rate function. This non-convexity of the large deviations rate function is associated to dynamical phase transitions in statistical mechanics [34, 35] and more recently in open quantum systems [12, 17, 26].

IV. EXAMPLES

Here we describe two examples illustrating the mathematical results. The first example is a quantum Markov chain where the large deviations rate function associated to the empirical measure and pair empirical measure shows dependence on some physical parameter. In the second example, the empirical measure rate function shows no dependence on a physical parameter, but the pair empirical measure rate function does. This example shows that, in order to uncover dynamical phase transitions through non-analyticities in a large deviations rate function, it may be necessary to consider higher-level statistics.

Example 1

Consider the two-dimensional quantum Markov chain with transition operator T acting on a density matrix $\rho \in M_2$ as

$$T(\rho) = V_0 \rho V_0^* + V_1 \rho V_1^* \quad (12)$$

where the Kraus operators are given by

$$V_0 = \begin{pmatrix} 0 & \delta \\ 0 & \epsilon \end{pmatrix}, \quad V_1 = \begin{pmatrix} \epsilon & 0 \\ \delta & 0 \end{pmatrix}$$

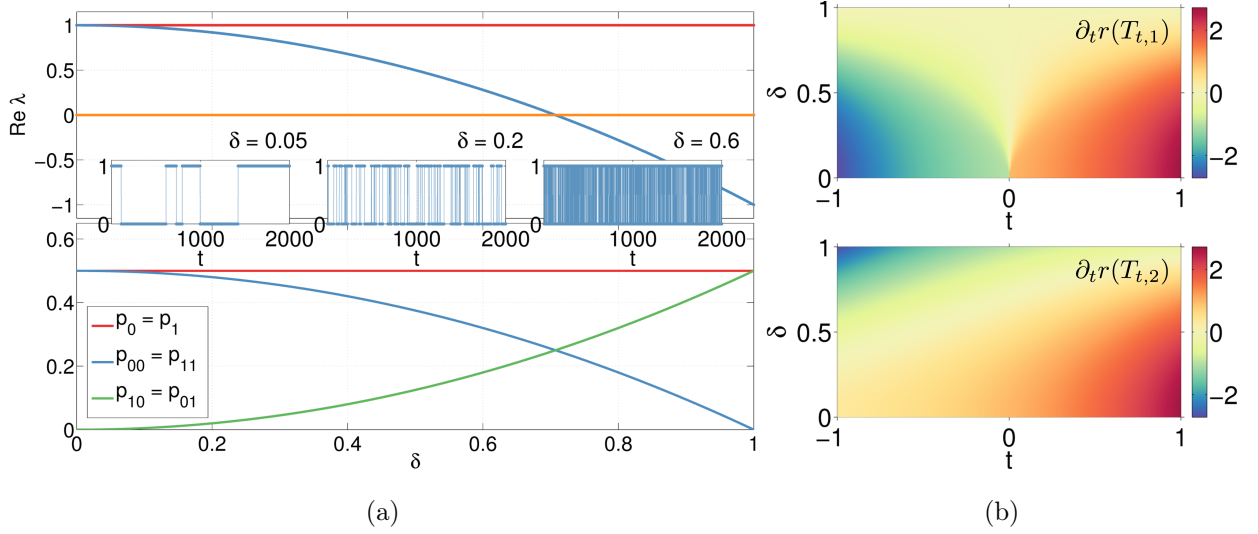


Figure 3: **(a):** Eigenvalues λ of T (top), statistics of trajectories (bottom): while the probabilities p_k are equal and constant, the pair probabilities $p_{i,j}$ vary with δ , which is reflected in the jump trajectories (see inset); **(b):** Derivatives of Sanov theorem level 1 and 2 spectral radii as functions of δ and the LD parameter with parametrisations $t[-1, 1] \in \mathbb{R}^2$ and $t[1, -1; -1, 1] \in M_2$, respectively. Note the discontinuity in $\partial_t r(T_{t,1})$ at $\delta = 0$ where the LDP no longer holds.

where $\epsilon = \sqrt{1 - \delta^2}$ with $0 \leq \delta \leq 1$. Defining $|u\rangle = \delta|0\rangle + \epsilon|1\rangle$ and $|d\rangle = \epsilon|0\rangle + \delta|1\rangle$, the Kraus operators can be expressed as $V_0 = |u\rangle\langle 1|$, $V_1 = |d\rangle\langle 0|$; the parameter δ interpolates between a trivial process at $\delta = 0$, where the Kraus operators project onto the eigenstates of σ_z , and the cyclic process with $V_0 = |0\rangle\langle 1| = V_1^*$ at $\delta = 1$. The eigenvalues of T are plotted in Fig. 3, showing that the set of eigenvalues reduces to $\{0, 1\}$ at $\delta = 0$ and $\{-1, 0, 1\}$ at $\delta = 1$.

Considering the level 1 and 2 statistics in the context of our Sanov theorem, in Fig. 3 we have plotted the probabilities p_k and $p_{i,j}$ to obtain $X^{(n)} = k$ and $(X^{(n)}, X^{(n+1)}) = (i, j)$, respectively, along the output trajectory in the stationary regime. By Thm. 3, the empirical measures associated to these jump statistics satisfy an LDP for $\delta > 0$, with rate function computed in Eq. (4) as the Legendre-Fenchel transformation of the spectral radius of $r(T_{t,m})$ of the associated transition operator. The first moment of the level k empirical measure is computed as the derivative $\partial_t r(T_{t,k})$ evaluated at $t = 0$; Fig. 3 shows the derivatives of the spectral radii $\partial_t r(T_{t,1})$ and $\partial_t r(T_{t,2})$. In this case, both the level 1 and 2 spectral radii

(and therefore the rate functions) show a dependence on δ , even though the first moment $\partial_t r(T_{t,1})|_{t=0}$ is constant.

Example 2

We now consider a quantum Markov chain which also satisfies Thm. 3, but where the level 1 statistics are independent of the physical parameter of the model. Let $\rho \in M_2$ be a density matrix and define the transition operator T as in Eq. (12), but where the Kraus operators are now defined as

$$V_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ i \sin \omega & \cos \omega \end{pmatrix}, \quad V_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos \omega & i \sin \omega \\ 0 & 1 \end{pmatrix}$$

where $0 \leq \omega \leq 2\pi$ (these dynamics may be obtained from a particular choice of parameters in a Heisenberg XYZ interaction between each noise atom and the system.)

Considering how the stationary states change with ω (see Fig. 4) we note that for $\omega = 0$, T becomes the identity map with full degeneracy of the eigenvalue 1. For $\omega \ll 1$ perturbation of the degenerate eigenvalue 1 shows that the eigenvalues split into $\lambda_1 = 1$, $\lambda_2 \approx 1 - \omega^2/2$ and a pair of complex conjugate eigenvalues $\lambda_3 \approx 1 + i\omega$, $\lambda_4 \approx \bar{\lambda}_3$. For $0 < \omega < \pi$ the stationary state ρ_{ss} is unique and a multiple of the identity, $\rho_{ss} = \frac{1}{2}\mathbf{1}$. At the point $\omega = \pi$ the Kraus operators are unitarily equivalent and stationary states are of the form $p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ with $0 \leq p \leq 1$.

As in the previous example, we consider the jump probabilities p_k and $p_{i,j}$ in the stationary regime (see Fig. 4). For $0 < \omega < \pi$ the probabilities $p_k, k = 0, 1$ are independent of ω with

$$p_k = \text{Tr}(V_k \rho_{ss} V_k^*) = \frac{1}{2} \text{Tr}(V_k^* V_k) = \frac{1}{2}$$

and similarly we obtain

$$\begin{aligned} p_{0,0} &= \frac{1}{4} (1 - \sin^2 \omega \cos \omega) = p_{1,1}, \\ p_{0,1} &= \frac{1}{4} (1 + \sin^2 \omega \cos \omega) = p_{1,0}; \end{aligned}$$

as shown in Fig. 4 this dependence on ω is reflected in the output trajectories, with increased intermittency when $p_{0,1} > p_{0,0}$.

By Thm. 3, the empirical measure associated to the level 1 and level 2 statistics on the output of this quantum Markov chain satisfies an LDP, with rate functions obtained from

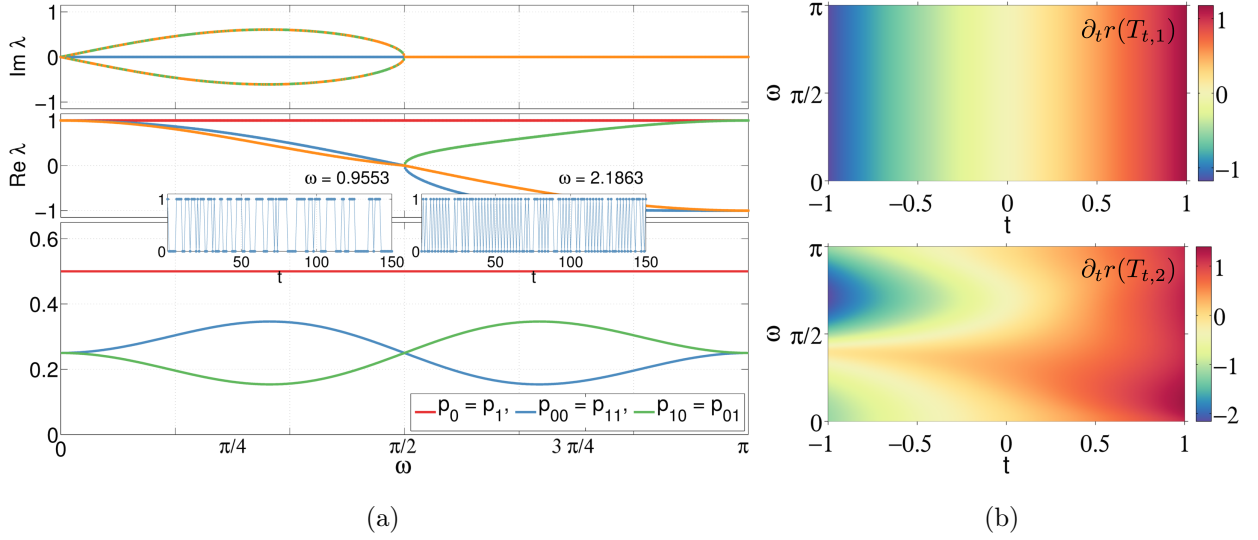


Figure 4: **(a)**: Eigenvalues λ of T (top), statistics of trajectories (bottom) and jump trajectories (inset); **(b)**: Derivatives of Sanov theorem level 1 and 2 spectral radii as functions of δ and the LD parameter with same parametrisation as in Fig. 3.

the corresponding spectral radii $r(T_{t,k})$. As shown in Fig. 4, the level 2 spectral radius $r(T_{t,2})$ depends on ω , while $r(T_{t,1})$ is constant, as we will now show.

Lemma 3. For the quantum Markov chain defined by T , the large deviations rate function associated to the level 1 empirical measure (i.e. sample mean) of the output process is independent of ω , for $0 < \omega < \pi$.

Proof. At the point $\omega = \pi/2$ the trajectories of the output process are equivalent to those of a classical process of i.i.d. fair coin tosses; this becomes evident by expressing the Kraus operators as

$$\begin{aligned} V_0 &= |u\rangle\langle 0|, \quad |u\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \\ V_1 &= |d\rangle\langle 1|, \quad |d\rangle = \frac{1}{\sqrt{2}}(i|0\rangle + |1\rangle) \end{aligned}$$

which always project onto the states $|u\rangle$ and $|d\rangle$ with equal probability. The large deviations rate function I associated to the sample mean of i.i.d. fair coin tosses is [32]

$$I(x) = \log 2 - x \log x - (1 - x) \log(1 - x), \quad 0 < x < 1.$$

By Thm. 3 the spectral $r(T_{t,1})$ is related to the rate function I by a Legendre transformation,

$$\begin{aligned}\log r(T_{t,1}) &= \sup_{0 < x < 1} \{xt - I(x)\} \\ &= t - \log 2 (1 + e^{-t})\end{aligned}$$

and so $r(T_{t,1}) = \frac{1}{2} (e^t + 1)$.

For $\omega \neq \pi/2$ it is easy to check that $\lambda_t = \frac{1}{2} (e^t + 1)$ remains an eigenvalue of $T_{t,1}$ with eigenmatrix

$$\rho_{ss} + \frac{1}{2} (\lambda_t - 1) \begin{pmatrix} 1 + \cos \omega & i \sin \omega \\ -i \sin \omega & 1 - \cos \omega \end{pmatrix}$$

where ρ_{ss} is the stationary state $\frac{1}{2}\mathbf{1}$. Since λ_t is independent of ω and $\lambda_t \rightarrow 1$ as $t \rightarrow 0$ we conclude that the moments $\partial_t^n \lambda_t|_{t=0}$ are independent of ω . \square

This example shows that the LD rate functions obtained from Thm. 3 are useful in uncovering dynamical behaviour of a system which is not immediately obtained from the lowest level LD picture.

V. DISCUSSION

We have shown that a large deviations principle holds for the empirical measure associated to an arbitrary number of subsequent outcomes obtained by measuring the output of a primitive quantum Markov chain. This extends the $m = 1$ large deviation result for the total counts of outcomes obtained in [27], which is the basis of the thermodynamics theory of quantum trajectories, and the theory of dynamical phase transitions [12, 17]. We presented an example in which the $m = 1$ LD rates are constant with respect to a system parameter, while the $m = 2$ theory captures this dependence. This suggests that a continuous-time version of our result would be relevant for a better understanding of dynamical phase transitions. Another direction in which the work can be extended is towards a Donsker-Varadhan LD theory for the empirical process of infinite trajectories.

Additionally, we showed that the empirical measure satisfies the Central Limit Theorem, extending the result from [20] which dealt with total counts statistics. The result, and its extensions to more general collective variables of the output are directly relevant for the statistical theory of system identification of open systems [21].

ACKNOWLEDGMENTS

Acknowledgments.— The authors would like to thank Juan Garrahan and Igor Lesanovsky for fruitful discussions. This work was supported by the EPSRC grant EP/J009776/1.

- [1] C. W. Gardiner and P. Zoller, *Quantum Noise*, 3rd ed., Springer Series in Synergetics (Springer, Berlin Heidelberg New York, 2004).
- [2] H.-p. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford New York, 2002).
- [3] B. Kümmerer, in *Coherent Evolution in Noisy Environments*, Lecture Notes in Physics, Vol. 611 (Springer, Berlin Heidelberg, 2002) Chap. 4, pp. 139–198.
- [4] M. Fannes, B. Nachtergaele, and R. F. Werner, *Communications in Mathematical Physics* **144**, 443 (1992).
- [5] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control* (Cambridge University Press, 2009).
- [6] G. Konya, D. Nagy, G. Szirmai, and P. Domokos, *Physical Review A* **86**, 013641 (2012).
- [7] S. Diehl, A. Tomadin, A. Micheli, R. Fazio, and P. Zoller, *Physical Review Letters* **105**, 1 (2010).
- [8] C. Ates, B. Olmos, J. P. Garrahan, and I. Lesanovsky, *Physical Review A* **85**, 043620 (2012).
- [9] B. Olmos, I. Lesanovsky, and J. P. Garrahan, *Physical Review Letters* **109**, 020403 (2012).
- [10] K. Netočný and F. Redig, *Journal of Statistical Physics* **117**, 521 (2004).
- [11] D. Andrieux, *Physics* **3**, 34 (2010).
- [12] J. P. Garrahan, A. D. Armour, and I. Lesanovsky, *Physical Review E* **84**, 21115 (2011).
- [13] T. E. Lee, H. Häffner, and M. C. Cross, *Physical Review Letters* **108**, 23602 (2012).
- [14] M. Foss-Feig, K. Hazzard, J. Bollinger, and A. Rey, *Physical Review A* **87**, 042101 (2013).
- [15] E. M. Kessler, G. Giedke, A. Imamoglu, S. F. Yelin, M. D. Lukin, and J. I. Cirac, *Physical Review A* **86**, 12116 (2012).
- [16] M. van Horssen and M. Guta, (2012), arXiv:1206.4956.
- [17] I. Lesanovsky, M. van Horssen, M. Guta, and J. P. Garrahan, *Physical Review Letters* **110**, 150401 (2013).

- [18] N. Malossi, M. M. Valado, S. Scotto, P. Huillery, P. Pillet, D. Ciampini, E. Arimondo, and O. Morsch, *Physical Review Letters* **113**, 023006 (2014).
- [19] C. Carr, R. Ritter, C. G. Wade, C. S. Adams, and K. J. Weatherill, *Physical Review Letters* **111**, 113901 (2013).
- [20] M. Guta, *Physical Review A* **83**, 062324 (2011).
- [21] M. Guta and J. Kiukas, *Commun. Math. Phys* (to appear) (2014).
- [22] O. Benson, G. Raithel, and H. Walther, *Physical Review Letters* **72**, 3506 (1994).
- [23] J. D. Cresser and S. Pickles, *Quantum and Semiclassical Optics: Journal of the European Optical Society Part B* **8**, 73 (1996).
- [24] J. P. Garrahan and I. Lesanovsky, *Physical Review Letters* **104**, 160601 (2010).
- [25] S. Genway, J. P. Garrahan, I. Lesanovsky, and A. D. Armour, *Physical Review E* **85**, 051122 (2012).
- [26] M. van Horssen, *Large deviations and dynamical phase transitions for quantum Markov processes*, Phd thesis, The University of Nottingham (2014).
- [27] F. Hiai, M. Mosonyi, and T. Ogawa, *Journal of Mathematical Physics* **48**, 123301 (2007).
- [28] I. Bjelaković, J.-d. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, and A. Szkořca, *Communications in Mathematical Physics* **260**, 659 (2005).
- [29] F. Hiai, M. Mosonyi, and T. Ogawa, *Journal of Mathematical Physics* **49**, 032112 (2008).
- [30] I. Bjelaković, J.-D. Deuschel, T. Krüger, R. Seiler, R. Siegmund-Schultze, and A. Szkořca, *Communications in Mathematical Physics* **279**, 559 (2008).
- [31] V. Jakšić, Y. Ogata, C.-A. Pillet, and R. Seiringer, *Reviews in Mathematical Physics* **24**, 1230002 (2012).
- [32] F. den Hollander, *Large Deviations*, Fields Institute Monographs (American Mathematical Society, Providence, Rhode Island, 2000).
- [33] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modeling and Applied Probability, Vol. 38 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010).
- [34] R. S. Ellis, *Scand. Actuarial. J.* , 97 (1995).
- [35] H. Touchette, *Physics Reports* **478**, 1 (2009).
- [36] W. Bryc, *Statistics & Probability Letters* **18**, 253 (1993).
- [37] S. Varadhan, *Large Deviations and Applications*, Regional Conference Series in Applied Mathematics (SIAM, Philadelphia, 1984).

- [38] J. Gärtner, *Theory of Probability & Its Applications* **22**, 24 (1977).
- [39] R. S. Ellis, *The Annals of Probability* **12**, 1 (1984).
- [40] I. Sanov, *Mat. Sb.* **42**, 11 (1957).
- [41] H. D. Miller, *The Annals of Mathematical Statistics* **32**, 1260 (1961).
- [42] M. D. Donsker and S. R. S. Varadhan, *Communications on Pure and Applied Mathematics* **29**, 389 (1976).
- [43] A. Buchleitner and K. Hornberger, *Coherent Evolution in Noisy Environments*, edited by A. Buchleitner and K. Hornberger, *Lecture Notes in Physics* (Springer-Verlag, Berlin Heidelberg New York, 2002).
- [44] S. Attal and Y. Pautrat, *Annales Henri Poincaré* **7**, 59 (2006).
- [45] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf, *Physical Review Letters* **95**, 1 (2005).
- [46] C. Schön, K. Hammerer, M. Wolf, J. Cirac, and E. Solano, *Physical Review A* **75**, 032311 (2007).
- [47] J. Dixmier, *Von Neumann Algebras* (North Holland, 1981).
- [48] D. E. Evans and R. Høegh Krohn, *Journal of the London Mathematical Society* **s2-17**, 345 (1978).
- [49] M. Sanz, D. Perez-Garcia, M. M. Wolf, and J. I. Cirac, *IEEE Transactions on Information Theory* **56**, 4668 (2010).
- [50] F. Fagnola and R. Pellicer, *Communications on Stochastic Analysis* **3**, 407 (2009).
- [51] T. Kato, *Perturbation Theory for Linear Operators*, 2nd ed. (Springer-Verlag, Berlin Heidelberg New York, 1976).